2019

Spatialized Auditory and Vibrotactile Cueing for Dynamic Three-Dimensional Visual Search

Rachel J. Cunio
Wright State University

SPATIALIZED AUDITORY AND VIBROTACTILE CUEING FOR DYNAMIC THREE-DIMENSIONAL VISUAL SEARCH

A thesis submitted in partial fulfillment of the

requirements for the degree of

Master of Science

by

RACHEL J. CUNIO

B.A., Saint Leo University, 2017

2019

Wright State University

APPROVAL SHEET (M.S.)

WRIGHT STATE UNIVERSITY

GRADUATE SCHOOL

APRIL 22, 2019

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Rachel J. Cunio ENTITLED Spatialized Auditory and Vibrotactile Cueing for Dynamic Three-Dimensional Visual Search BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science.

Joseph Houpt, Ph.D.
Thesis Director

Scott Watamaniuk, Ph.D.
Graduate Program Director

Debra Steele-Johnson, Ph.D.
Chair, Department of Psychology

Committee on
Final Examination

Brian Simpson, Ph.D.

J. Christopher Brill, Ph.D.

Scott Watamaniuk, Ph.D.

Barry Milligan, Ph.D.
Interim Dean, Graduate School

ABSTRACT

Cunio, Rachel J. M.S. Department of Psychology, Wright State University, 2019.
Spatialized Auditory and Vibrotactile Cueing for Dynamic Three-Dimensional Visual Search.

The traditional method of maintaining spatial awareness through visual displays can cause visual system overload and lead to performance decrements. This study examined the benefits of spatialized auditory, tactile, and audio-tactile cues for maintaining awareness as a method of enhancing visual search performance. I examined visual search performance in an immersive, dynamic, three-dimensional (360-degree), virtual reality environment with no cues, spatialized auditory cues, degraded spatialized auditory cues, spatialized tactile cues, spatialized audio-tactile cues, and degraded spatialized auditory with tactile cues. Results indicated a significant reduction in visual search time from the no-cue condition when any cues were presented. The tactile display did not provide an additional benefit when combined with the auditory display. The results of this study can inform the creation of multimodal displays appropriate to specific operational settings, such as including auditory displays in dynamic settings or including tactile displays when the visual target is behind the operator, which will improve visual search performance, increase mission effectiveness, and possibly save lives.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First, I would like to thank my advisor, Dr. Joseph Houpt for his research insights and encouragement. I would also like to thank my committee members, Dr. J. Christopher Brill, Dr. Brian Simpson, and Dr. Scott Watamaniuk for their support and expertise. I have extra gratitude for Chris and Brian, as well as the Battlespace Acoustics Branch of the Air Force Research Laboratory at Wright-Patterson Air Force Base for their support with equipment, lab space, and research participants on this project. Thanks also to Dr. Paul Havig for introducing me to multimodal research during the first month of my graduate career and supporting me with opportunities since then, and to David Dommett, for his programming expertise and countless hours of equipment and software testing.

I am grateful for my labmates Cara Zinn, Lori Mahoney, Joey Glavan, Betsy Fox, Hanshu Zhang, Sam Johnston, and Jane Hwang. It is a joy to work with such talented people and I hope to have the opportunity to work with many of them in the future. Another big thank-you goes out to the friends who have supported my efforts and sleepless nights for the past two years, especially those from the Wright State University Catholic Campus Ministry. They have inspired me to keep going and not lose faith even when the stress seemed overwhelming. The final and most important expression of gratitude is for my family, especially my parents, Stephen and Paula Cunio, for their never-ending love and support.

I. INTRODUCTION

Many operational settings require visual searches of spatial environments. Traditional methods of presenting spatial information to aircraft operators involve heads-up displays (HUDs), which communicate altitude, direction, distance, and other spatial information visually. However, HUDs can overload the visual system of the operator, causing important information to be missed and negatively affecting the operator's visual search effectiveness, even in static environments (e.g., Stokes & Wickens, 1989). In an operational or military setting, decreased speed and/or accuracy of visual search can lead to mission failure, or even fatality. To combat visual system overload, researchers have examined the utility of presenting spatial information via multimodal cueing, particularly spatialized auditory and/or tactile displays (e.g., Brill, Lawson, Rupert, & Gagliano, 2015; Simpson, Brungart, Dallman, Joffrion, Presnar, & Gilkey, 2005). Whereas researchers have conducted studies in two-dimensional dynamic environments (e.g., McIntire et al., 2010) or three-dimensional static environments (e.g., Mateo, Simpson, Gilkey, Iyer, & Brungart, 2012), they have not examined the benefits of display types in three-dimensional (360 degree) and dynamic (moving) environments. And yet, dimensionality and movement both are fundamental components of operational visual search applications, such as pilots searching for other airborne aircraft. Additionally, researchers rarely consider a three-dimensional (3D), dynamic context when examining the cognitive processes involved in visual search. Thus, the purpose of my research was to address how dynamic visual search performance in a 3D virtual environment changes when spatialized auditory and/or vibrotactile cues are given, as compared to when no cues, or degraded cues are given.

Visual Search

Understanding the limitations of the human visual system is important when addressing visual search. Whereas humans have excellent visual acuity within the central visual field (the fovea), the periphery of the visual system has relatively low spatial resolution. Additionally, the visual field spans only about 135 degrees vertically and 200 degrees horizontally (Wandell, 1995). This means that the human visual system's field of view only contains about half of the spatial world at any particular point in time. Even within the field of view, attention is also required for effective spatial awareness (Endsley, 2000). Therefore, while the visual system is useful for signaling within the field of view, operational settings require omnidirectional spatial awareness. Fortunately, the auditory and tactile modalities have practical and perceptual benefits that can supplement the limitations of the visual system (e.g., Brill et al., 2015). Therefore, it is important to explore the benefits of auditory and tactile cueing in an environment that reflects the nature of operational settings in which future multimodal displays might be employed.

Because of its ubiquitous nature, researchers have studied visual search for many years (e.g., Kingsley, 1932). However, from the debut of research on visual search with mounted pictures on a cardboard display (Kingsley, 1932) to its more recent presentation on desktop displays (e.g., Amor et al., 2016; Wolfe, 1998), researchers have studied visual search primarily in the context of static environments, particularly two-dimensional situations. Also, mean response time (RT) is often the primary variable of interest. However, traditional mean RT analyses do not capture all assets of the cognitive processes involved in visual search, which are more informative for the creation of multimodal displays for operational settings.

Understanding the characteristics of the human visual system has been useful in examining static visual search performance, but researchers have not addressed the complexity of the processes of the human visual system in its entirety, particularly in dynamic and 3D environments that require high amounts of eye and head movement. This points to a need for more extensive research into the cognitive processing involved in visual search, particularly in dynamic, 3D environments.

Auditory Cueing

Because of the aforementioned limitations of the visual system, researchers often have provided cues through other modalities for visual search tasks (e.g., Santangelo & Spence, 2007; Stein, Wallace, & Stanford, 2001). One modality that researchers have used to provide cues in visual search tasks is audition. Whereas the auditory system has much lower acuity than the visual system, the omnidirectional nature of the auditory system makes spatialized auditory cues particularly beneficial for cueing events in 3D spaces. In fact, previous research has shown that spatialized auditory cues can effectively decrease visual search time in 3D space as compared to non-cued conditions (Bolia, D'Angelo, & McKinley, 1999; Perrot, Cisneros, McKinley, & D'Angelo, 1996) and nonspatialized auditory cues (Flanagan, McAnally, Martin, Meehan, & Oldfield, 1998). Also, Simpson et al. (2005) showed the utility of spatial auditory cues for providing directional information.

Researchers have shown that virtual spatial auditory displays effectively communicate spatial information in dynamic (moving) visual search tasks (McIntire, Havig, Watamaniuk, & Gilkey, 2010). Some researchers have asserted that the decrease in response time found when spatialized auditory displays are used arises from the fact that auditory cues solicit a reflexive orienting response that effectively decreases the search space required for the visual system to complete (Brill, Lawson, & Rupert, 2014). Whereas this assertion might hold true, I wanted to explore deeper cognitive benefits by showing the utility of spatialized auditory cueing beyond simply reducing search space. Whereas some studies have used auditory displays to portray both 3D (spatialized) information and information about moving targets (e.g., McIntire et al., 2010), most studies that involve dynamic stimuli have used simple orienting auditory cues, rather than continuously updated spatialized cues (e.g., Mateo et al., 2012). So, whereas previous research has found that cueing of various types can be beneficial to visual search, implementing a cue type in this study that more accurately reflects the nature of the stimuli (dynamic and spatialized) allowed for greater insight into the cognitive processing resources available during a cued visual search task.

Whereas the auditory system is beneficial for supplementing visual information in omnidirectional environments, there are a number of limitations to the utility of spatial auditory displays. For example, the presence of noise has been shown to degrade sound localization accuracy (Good & Gilkey, 1996). Spatial ambiguity such as front-back reversals can also lead to a loss of effective spatial awareness (e.g., Mateo et al., 2012). Therefore, in addition to examining the benefits of auditory enhancement of visual search, I wanted to examine a modality to supplement auditory information during visual search.

Tactile Cueing

Another way to present spatial information that can complement the weaknesses of the auditory system in adverse auditory environments way is through the tactile modality. Researchers have shown that tactile cues communicate spatial information effectively, especially when visual or auditory modalities are compromised (e.g., Brill & Scerra, 2014; Cholewiak, Brill, & Schwab, 2004; Rupert, 2000). In fact, Brill et al. (2015) reported that tactile cues can disambiguate front-back reversals. Also, Brill, Lawson, and Rupert (2014) and Van Erp et al. (2006) have demonstrated how vibrotactile displays can effectively reduce target search times and signal spatial orientation information and directional threats, considerations which are particularly applicable to visual search tasks. Further, Mateo et al. (2012) found that vibrotactile cues can be effective in omnidirectional (3D) environments. This means that the tactile modality provides the ideal complement to the audition modality for spatial localization.

The tactile system's acuity in for spatial information varies based on the density of the receptors on the particular area of the body, but it can be as small as four millimeters (Sherrick & Cholewiak, 1986). There are many types of tactile displays, but most recent studies have used vibrotactile stimulators placed around the torso area (e.g., Haggit, 2014; Hancock et al., 2013; Mateo et al., 2012). Acuity for vibrotactile stimulation is much lower and more difficult to define (Brill, 2015). In similar fashion to auditory cueing research, tactile cues are often presented as brief vibration bursts intended to solicit an initial orienting response, rather than continuously updated vibrations to provide continuously updated spatial information. Continuously updated cues are more effective in dynamic environments because the spatial location of the visual target is constantly changing. I wanted to increase the congruency between the display type and the visual stimuli in the dynamic visual search task so as to allow for greater insight into the underlying cognitive mechanisms of visual search. Therefore, I sought to use a tactile display that provided continuously updated spatial information, rather than simple vibration bursts.

Multisensory Cueing

Many studies have demonstrated that multimodal cueing, specifically auditory and vibrotactile combined displays, can facilitate localization of targets or at least direct attention to cued locations in the environment (Hancock et al. 2013; Santangelo & Spence, 2007). The auditory and tactile systems have complementary limitations and benefits. For example, auditory localization is highly accurate along the left/right dimension, but sounds that produce the same interaural time and level differences are difficult to localize. This is especially true for frontal and rear localization by the auditory system, making fore-aft reversals common in auditory localization situations. In contrast, the tactile system is highly accurate for fore-aft localization but can be inaccurate for properly identifying target elevation when targets or cues are presented in lateral positions (Brill et al., 2015). Additionally, the auditory system receives information distally, whereas vibrotactile cues must be proximal (i.e., have contact with the body). Clearly, the intrinsic qualities of the auditory and tactile systems are complementary by nature.

Despite the complementary nature of the auditory and tactile sensory systems, display types or apparatuses constraints might limit the effectiveness of multisensory cues. Mateo et al. (2012) found no significant advantage of audio-tactile cueing over auditory only displays while using free-field spatialized auditory cues. The inverse-effectiveness principle (Stein, Wallace, & Stanford, 2001) suggests that this may be due to the highly effective auditory cue used in the experiment. One method of examining the nature of this sensory systems is to match the effectiveness of the cues. Haggit (2014) found that degrading the spatial information in an auditory cue led to an audio-tactile display advantage over auditory display alone in a 3D environment. However, degrading spatial information is not as true to the nature of operational environments. In an operational environment, such as a cockpit, auditory noise, rather than degraded spatial information, leads to degraded auditory environments. One reason for including vibrotactile cues in addition to auditory cues is to provide redundant cueing information when the auditory environment is noisy (Brill et al., 2014). For an experimental situation in which the auditory cue is highly effective, simulating a degraded auditory environment might impose greater dependency on tactile cues. Therefore, I included a degraded (noisy) auditory cueing condition to determine whether Haggit's (2014) findings in a degraded auditory environment were robust to a noisy environment with dynamic stimuli.

In general, multimodal studies that emphasize cross-modal matching have shown higher benefits of multimodal information displays. Whereas these studies have addressed the benefits of multisensory cueing in 3D environments, they have failed to consider multisensory cueing in the context of dynamic situations.

Analytic Approach

Whereas implementing a cue type that provided continuously updated information about targets was more true to the nature of the stimuli (dynamic and omnidirectional), I wanted to also address the complexity of processes in the human visual system in their entirety. My analytic approach included traditional mean RT analyses because they are widely recognized by the scientific community. However, using only mean RT analyses poses two major problems. One problem is that they fail to consider the full range of RTs, meaning that potentially meaningful RT distribution shape information is absent in the analyses. Another problem is that they involve no psychologically meaningful baseline for comparison of RTs. This means that traditional mean RT analyses allow only for a dichotomous categorical distinction of "better" or "worse" when comparing conditions to each other, which is not sufficient when examining cognitive processing. Therefore, my analytical approach included workload capacity analyses to examine the entire distribution of RTs and to use a meaningful baseline to examine the underlying cognitive processes involved in the experimental visual search task, beyond the limits of traditional mean response time analysis.

Workload capacity analysis uses the Systems Factorial Technology (SFT) framework to determine how the processing speed of individual sources of information is affected by the inclusion of other sources of information (Houpt, Blaha, McIntire, Havig, & Townsend, 2013). This framework has demonstrated utility for understanding cognitive processes in previous multimodal analyses (e.g., Fox, Glavan, & Houpt, 2014). For this study, the visual, auditory, and tactile scenes were each treated as an information source, and processing capacity was examined in terms of response time using capacity coefficients, a measure of workload capacity in the SFT framework.

First, to examine the full range of response times, I wanted to compare each cueing condition to the condition involving only the visual task, with no additional cues. The structure of the experimental task requires the single-target self-terminating ("ST-ST") capacity coefficient, meaning that the first detection of the target is used to make the response. The "ST-ST" capacity coefficient is the ratio of cumulative reverse hazard function (a running measure of work completed for specific time) of RTs for processing the target, V (the visual modality) in isolation, $K_V(t)$, to the cumulative reverse hazard function of RTs for processing the target within a context of other sources of information, $VM(t)$ (the visual modality and modality of the cue type combined), $K_{VM}(t)$. In this case, the no-cue condition is considered to be the visual modality, and any cued condition is a combination of the visual modality and the modality through which the particular cue is provided (i.e., auditory, A, tactile, T, or audio-tactile, AT). Then, this ratio of observed performance is compared to a baseline performance. To provide a psychologically meaningful baseline, workload capacity analysis assumes unlimited capacity, independent, parallel (UCIP) processing of information sources. The "ST-ST" capacity coefficient for the comparison of the no-cue condition to the auditory condition is represented as an example by the following equation:

$$C_{STST}(t) = \frac{K_V(t)}{K_{VA}(t)}. \quad (1)$$

The “ST-ST” capacity coefficient characterizes processing as limited, unlimited, or super capacity. Limited capacity, when $C_{STST}(t) < 1$, refers to the capacity of a system in which performance decreases as more sources are added (an increased workload), indicating worse performance than the baseline model. Unlimited capacity, when $C_{STST}(t) = 1$, refers to the capacity of a system in which performance is consistent with the baseline as more sources of information are added, indicating performance equal to the baseline model. Super capacity, when $C_{STST}(t) > 1$, refers to the capacity of a system in which performance increases as more sources of information are added, indicating performance better than the baseline model.

“ST-ST” capacity coefficients are useful for making a categorical distinction of “better” or “worse” across the whole range of RTs when comparing each cueing condition to the condition involving only the visual modality. However, when I compared the audio-tactile condition to the audio and tactile conditions respectively, I needed to go beyond categorization to understand how much better or worse participants were compared to the unlimited processing capacity that the baseline assumes. To do this, I needed to use another type of capacity coefficient, considering how many information sources are processed before a participant can make a decision in the experimental task.

The capacity “OR” processing structure describes a task in which participants process all information sources simultaneously (i.e., in parallel), but use the first information source for which they complete processing to make a decision. This processing is characterized as “OR” processing because only one information source or another is required to complete processing, as opposed to both or all sources being required to finish being processed before the participant can make a response. For my task, participants were asked to locate a visual target among distractors and make a visual discrimination about the target. Sometimes they were provided with auditory, tactile, or audio-tactile cues to assist in the localization of the target. Participants could use any of the modalities provided during the particular condition to localize the visual target. Whichever cueing modality for which the participant finished processing first allowed the participant to locate the visual target (but not make the visual discrimination) make a response (based on the visual discrimination following localization). For example, in the audio-tactile condition, the participant could have processed the auditory display, the tactile display, and the visual information that is inherent in the task. If the auditory information processing was completed first for localization, then the participant could have moved into the visual discrimination part of the task, as demonstrated in Figure 1 below. Therefore, the structure of this experimental task was characterized as “OR” processing, and the capacity “OR” coefficient analysis was ideal for comparing the multimodal cueing conditions to the single-modality cueing conditions.

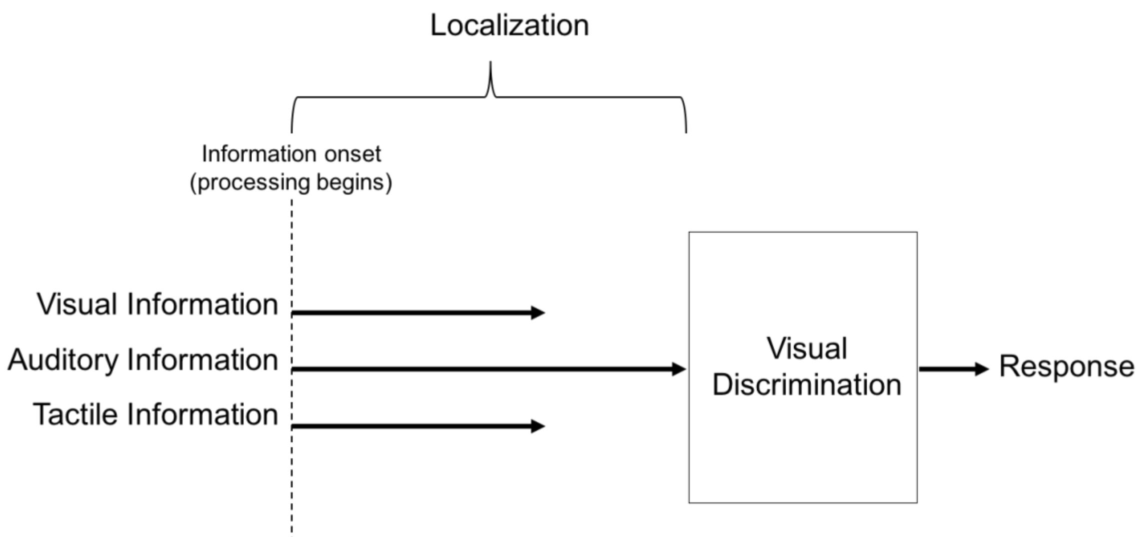

Figure 1. Visual representation of information "OR" processing in the experimental task. Each arrow represents the processing speed for each information source. The dashed line represents the point in time when the information sources are presented to the participant and processing begins. The figure shows the participant completing processing of auditory information first, allowing the participant to move to the decision portion of the task and then make a response.

Further, I needed to consider the concept of statistical facilitation, a phenomenon that makes using the UCIP baseline model for meaningful comparison particularly important. In redundant cueing situations, such as may be the case in some multimodal cueing conditions, reductions in RT occur because perceptual processes can combine when the same information is presented through multiple sources (e.g., through the auditory and tactile modalities) with the goal of making one response (e.g., indicating the location of a target). Assuming the response times produced from processing each modality are independent random variables, the fastest response time will be faster than either modality alone, regardless of whether the presence of both modalities aided in the processing of either modality individually. This phenomenon, causing RT speed-ups, is referred to as statistical facilitation. Using the UCIP model as a baseline allows for comparison of observed data to predictions based on a model that assumes statistical facilitation will occur. This means if the results show something other than the baseline prediction, the UCIP model can be rejected and there may be either statistical debilitation occurring if the observed data are worse than the baseline (see Townsend & Wenger, 2004), or cognitive facilitation if the observed data are better than the baseline. The most desirable multimodal cueing situation would involve cognitive facilitation, which would show that individual source processing speed increases with the addition of more information sources (or in this case, modalities).

The capacity "OR" analysis provides a metric for understanding the extent to which my participants performed better or worse than statistical facilitation would predict. Similar to capacity "ST-ST" coefficients, capacity "OR" coefficients characterize performance as limited, unlimited, or super capacity, depending on performance relative to the baseline. However, the capacity "OR" characterization relates limited capacity to statistical debilitation, unlimited capacity to statistical facilitation, and super capacity to cognitive facilitation.

The capacity "OR" coefficient is the ratio of cumulative hazard function (a running measure of work completed for specific time) of response times to redundant signals at time t for visual signals, V, auditory signals, A, and tactile signals, T. This is appropriate for a comparison of audio-tactile conditions to auditory and tactile conditions because the visual modality is present for all conditions, and therefore I assumed that the visual modality was as helpful in the multimodal condition as it was in any single-modality conditions. The UCIP baseline model predicted that the sum of the cumulative hazard functions (running measures of workload at time, t) for the visual modality, H_V, the auditory modality, H_A, and the tactile modality, H_T, should equal the cumulative hazard function for the combination of the three modalities, H_{VAT}, as shown in Equation 2 below:

$$H_{VAT}(t) = H_V(t) + H_A(t) + H_T(t). \tag{2}$$

The standard capacity "OR" analysis requires comparing each of the single modality conditions (i.e., visual, auditory, or tactile) to the condition with all modalities (i.e., visual, auditory, and tactile combined). However, the nature of my experiment as a visual search task required that the visual information be present in all conditions. This means that my experiment did not involve a condition that included auditory but not visual information, nor a condition that included tactile but not visual information, but there was a condition that involved both auditory and visual information and a condition that involved both tactile and visual information. Fortunately, the UCIP baseline model predicted that the cumulative hazard function at time, t, of the visual and auditory information together, H_{VA}, equals the sum of the cumulative hazard functions of visual information alone, H_V, and auditory information alone, H_A, yielding the following equation:

$$H_{VA}(t) = H_V(t) + H_A(t). \tag{3}$$

Which, solved for H_A, reads:

$$H_A(t) = H_{VA}(t) - H_V(t). \tag{4}$$

The same assumption and rearrangement hold for visual and tactile information, in Equations 5 and 6 below:

$$H_{VT}(t) = H_V(t) + H_T(t) \tag{5}$$

$$H_T(t) = H_{VT}(t) - H_V(t). \tag{6}$$

Therefore, I can substitute the right-hand side of Equation 4 for $H_A(t)$ in Equation 2 and the right-hand side of Equation 6 for $H_T(t)$ in Equation 2, and solve as follows:

$$\begin{aligned} \mathrm{H}_{VAT}(t) &= H_V(t) + H_A(t) + H_T(t) \\ &= \mathrm{H}_V(t) + H_{VA}(t) - H_V(t) + H_{VT}(t) - H_V(t) \\ &= \mathrm{H}_{VA}(t) + H_{VT}(t) - H_V(t) \end{aligned}. \tag{7}$$

The final line of Equation 7 in ratio form can be compared to the UCIP baseline model to yield capacity "OR" coefficients in the form of the following equation:

$$C_{OR}(t) = \frac{H_{VAT}(t)}{H_{VA}(t) + H_{VT}(t) - H_V(t)}. \tag{8}$$

The "OR" capacity coefficient characterizes processing as limited, unlimited, or super capacity. Limited capacity, when $C_{OR}(t) < 1$, refers to the capacity of a system in which performance decreases as more sources are added (an increased workload), indicating statistical debilitation in processing speed. Unlimited capacity, when $C_{OR}(t) = 1$, refers to the capacity of a system in which performance is consistent with the baseline as more sources of information are added, indicating statistical facilitation in processing speed. Super capacity, when $C_{OR}(t) > 1$, refers to the capacity of a system in which performance increases as more sources of information are added, indicating cognitive facilitation in processing speed. Therefore, capacity "OR" provides an appropriate metric for understanding how much better or worse than statistical facilitation (unlimited capacity) participants were in conditions involving multimodal cues (i.e., audio-tactile cues) compared to conditions involving single-modality cues (i.e., auditory or tactile).

In summary, visual search, an important task in operational settings, is limited by the problem of visual system overload. Research has shown the utility of virtual spatial auditory displays for aiding dynamic visual searches in 2D space (e.g., McIntire et al., 2010) and static visual searches in 3D space (e.g., Mateo et. al, 2012). Further, research has shown the utility of tactile cues in aiding spatial awareness in dynamic environments (e.g., Van Erp et al., 2006). However, many visual search tasks that occur in operational environments involve both moving operators or targets and real-world omnidirectional targets and distractors. Therefore, a logical extension of the previous research was to examine how spatial auditory cueing can facilitate (or inhibit) performance on a dynamic visual search task in a 3D environment. Additionally, examining the entire range of RT and examining the cognitive processes involved in visual search can inform the specific methods in which future multimodal displays to assist visual search should be created. Thus, the primary goal of my research was to evaluate auditory and tactile enhancement of visual search. To do this, I expanded the current visual search research in two ways. First, I provided spatialized auditory, tactile, and audio-tactile cues for a visual search task in a 3D, dynamic environment. Second, I examined the cognitive processes involved in visual search using novel statistical techniques.

Hypotheses

Accuracy. Accuracy information arose from the responses regarding the visual discrimination. This visual discrimination required only visual information, which was inherent in the task and therefore present for all cueing conditions. Therefore, I expected that participants would show high accuracy overall.

Mean response time. Because Mean RT analysis is well recognized by the scientific community and I wanted to be able to compare my results to previous studies that used mean RT analysis, I started with traditional mean RT analysis. I expected that cueing of any form would lead to significant reductions in visual search times over the no-cue condition, but there would be differences in mean RT across cueing conditions. Specifically, I predicted that the conditions involving degraded cues would provide less of an RT speed-up than the non-degraded conditions.

Workload capacity. I also hypothesized that when examining processing capacity using the "ST-ST" capacity coefficient, participants would show super capacity processing across the entire distribution of RTs, meaning that any RT speed-up occurring from the cues over the no-cue condition would extend across the entire range of RTs. I also expected that when examining processing capacity using the "OR" capacity coefficient, participants would show unlimited or super capacity processing, meaning that the addition of a second cue type in the multimodal cueing conditions would provide an additional RT benefit over the single-modality conditions.

II. METHOD

Participants

Fourteen participants (seven male) between the ages of 18 and 34 (M = 23.9, SD = 3.9) from the Dayton, Ohio area, including 10 participants from the research subject panel at the Air Force Research Laboratory at Wright-Patterson Air Force Base (AFRL/WPAFB), completed the experiment. A power analysis considering the assumed statistical power of 0.8 and the expected Cohen's d effect size of 1.7, derived from the results of a previous dynamic visual search study (McIntire et al., 2010), determined the number of required subjects (14). All subjects were screened via standard processes for normal hearing and normal or corrected-to-normal vision.

Participants completed three experimental sessions, each lasting approximately one hour with at least one hour of rest between sessions and no more than two sessions per day. Session 1 included Part 1 and Part 2 of training as well as two blocks of experimental trials. Sessions 2 and 3 each contained four blocks of experimental trials. AFRL/WPAFB research panel subjects were compensated for their time based on their normal hourly pay rate of $10 to $15 per hour. The four remaining subjects were compensated with Visa gift cards at a rate of $15 per hour. Participants who took longer than one hour to complete an experimental session were compensated for two hours.

Design

The experiment employed a within-subjects design in which participants were asked to complete a visual search task (i.e., locate a target among distractors). During the experiment, spatialized auditory and/or spatialized tactile displays were presented in various forms (i.e., absent, present, or degraded), as shown in Table 1 below, to assist participants with localization of the target. I presented all participants with six cueing conditions: no cue, auditory cue, degraded auditory cue, tactile cue, auditory with tactile cue (i.e., audio-tactile cue), and degraded auditory with tactile cue.

Table 1

Experimental Trial Conditions

	Auditory Cue		
Tactile Cue	Absent	Present	Degraded
Absent	No-Cue	Auditory Cue	Degraded Auditory Cue
Present	Tactile Cue	Audio-Tactile Cue	Degraded Auditory with Tactile Cue

Visual search task. The visual search task for this study involved localizing a Landolt C target stimulus (a ring with a specified gap in a particular location) among 100 distractor rings and reporting the location of the gap (i.e., left or right). The visual target was identical to the distractors except the target included a small gap on either the left or right side of the ring. This design ensured that participants were both detecting the presence of the target and discriminating its gap location.

This distinction is important because for analysis purposes, the experimental task can essentially be broken into two parts: localization and discrimination. Localization involved locating the Landolt C. Discrimination was the decision portion that required determining the location of the gap on the visual target. Localization can involve any of the cues (i.e., visual, auditory, and/or tactile) provided to the participant for the given trial. After the visual target is located, its gap discrimination only involves the visual modality because the auditory and tactile cues provide only spatial information, not target gap information. Therefore, there was no reason to expect any accuracy differences across conditions if the assumption that the visual modality was the only modality involved in target discrimination holds. Additionally, I considered all RT analyses to examine the localization portion of the task, because RT for discrimination should be consistent across conditions once the visual target is localized.

Participants responded by indicating the location of the gap on the target using left and right triggers, respectively, on wireless Oculus Rift virtual reality system hand controllers. I measured RT from the beginning of trial onset until the participant ended the trial by responding. I collected accuracy information (i.e., correct or incorrect) although I did not expect to see differences in accuracy across conditions. Additionally, I collected streaming participant head-position and target location position data for use in a future project. Room-scale head tracking allowed for adjustment of visual, auditory, and tactile stimuli based on spatial position of the participant.

Visual stimuli. Previous studies using spatialized auditory and audio-tactile cues have been conducted in the Auditory Localization Facility (ALF) of the AFRL/WPAFB, which is a physical geodesic sphere suspended within an anechoic chamber, used for conducting research on spatial audition (e.g., Haggit, 2014; Mateo et al., 2012).

In order to ensure comparability of outcomes between physical- and virtual-worlds, I created the experimental virtual space to contain a sphere with the same radius (2.3 m) as ALF. I presented the visual stimuli through a corded Oculus Rift virtual reality system that produced a 2160 x 1200 resolution image with a refresh rate of 90-Hz. Participants only had rotational body movement capabilities (although head movement was not restricted), so the radius of the virtual sphere was fixed throughout the experiment. The experiment was programmed using *Unity*, a cross-platform video game engine. The application for running the experiment is available upon request. The visual distractors and targets were black rings and Landolt C's, respectively, moving along the inner surface of a gray sphere. Adapting the same visual stimuli used by McIntire et al. (2010), the width of each visual stimulus was 0.96 degrees of visual angle, and the gap size on the Landolt C visual targets was 0.12 degree of visual angle. These dimensions required participants to foveate within five degrees of visual angle of the visual target in order to discriminate its category (i.e., left or right; McIntire et al., 2010). Figure 2 below shows an example visual target and an example visual distractor.

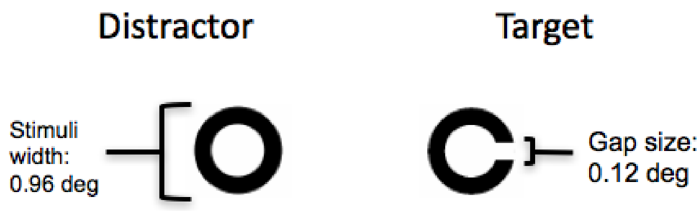

Figure 2. Visual distractor and target. The width (diameter) of both the distractor (left) and the target (right) was 0.96 degrees of visual angle. The gap size of the target was 0.12 degrees of visual angle.

It was possible for the stimuli to move 360 degrees of azimuth (rotation around the participant), but only 90 degrees of elevation (i.e., +/- 45 degrees from the participants' head position; see Figure 3). Visual stimuli moved at 10 degrees of visual angle per second, matching previous dynamic visual search research (McIntire et al., 2010). The density of distractors relative to search field dimensions also matched McIntire et al. (2010), with each trial containing one target and 100 distractors. Trajectories consisted of orbits around the participant along the inner surface of the sphere, as shown in Figure 3. I calculated starting points and trajectories of targets and distractors such that stimuli themselves were never closer than the size of one visual stimulus (i.e., 0.96 degrees of visual angle) although the trajectory paths themselves could cross.

Figure 3. Experimental virtual environment. The virtual sphere allowed visual stimuli to move anywhere in the dark blue shaded region, (i.e., 360 degrees of azimuth (rotation around the participant) and +/- 45 degrees from the participants' head position) along the inside of the sphere, at a constant speed of 10 degrees visual angle per second.

Auditory stimuli. The auditory display included virtual spatialized sounds corresponding to the location of the dynamic visual targets (using the *slab3d* spatial audio rendering software; http://slab3d.sonisphere.com). This means that the auditory signals presented to the participant sounded as though they were generated from the current visual target location. For example, if the current location of the visual target was to the right of the participant, the target cue sounded to the participant as if it were coming from the participant's right. The auditory target cue consisted of three 100-Hz click train pulses, each 100-ms long, with a 50-ms silent interval between pulses one and two and pulses two and three. After the third pulse, there was a 500-ms silent interval. This cue was chosen based on its utility as demonstrated in previous auditory localization studies (e.g., Brungart & Simpson, 2008). This temporal profile was chosen for the auditory cue based on studies that presented similar auditory cueing situations (e.g., McIntire et al, 2014).

I presented the auditory stimuli as 'present' (i.e., not degraded) or 'degraded', (depending on the trial) in order to examine the effects of degrading the auditory environment on visual search performance. In the 'present' auditory condition, only the target auditory stimulus was presented. In the 'degraded' auditory condition, in addition to the target stimulus, diffuse broadband white noise was presented as an auditory masker to create a degraded auditory environment. The diffuse noise source was a virtual location directly below the participant's feet. Both the target and the masker were presented at the comfortable listening level of 78 dB SPL and controlled by a generic set of head-related transfer functions (HRTFs), displayed through corded Sennheiser HD 280 professional headphones. The sound level of both the target and the masker were determined during pilot studies based on particular equipment requirements, as specified in the tactile stimuli section below. The target and/or masker (depending on the condition) were presented on continuous loops of 900-ms and 10-s, respectively, until the participant responded, ending that particular trial.

Tactile stimuli. The tactile display included spatialized signals corresponding to the visual target trajectories. I presented the tactile stimuli through eight vibrating tactors attached to an elastic belt with a Velcro fastener. I used standardized measurement procedures to ensure uniform placement of tactors across subjects and within subjects across trials. The tactile belt was worn around the participants' abdomen with the stimulators positioned approximately 1 inch above the navel. Each tactor was responsible for a 45-degree slice of the sphere (i.e., +/- 22.5 degrees left and right of the tactor), as shown in Figure 4.

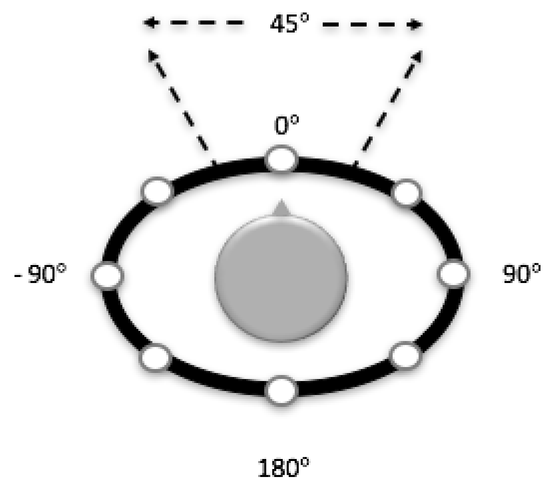

Figure 4. Top-down view diagram of vibrotactile display set-up. The gray circle in the middle represents participant head position, with the gray triangle indicating the direction of the nose. The black ring around the participant head represents the elastic belt on which the tactors were attached. The small white circles indicate the locations of each tactor. The dashed arrows indicate the slice of the virtual sphere for which the tactor could present spatial information.

I presented the tactile stimuli through Engineering Acoustics Incorporated (EAI) C2 linear transducer tactors that vibrated much like a vibrating cell phone against the skin and were controlled using an EAI Tactor Control Unit. The tactors' frequency response was 250-Hz, and the tactors were attached to a corded 2-inch wide elastic belt with a

Velcro fastener. I had two sizes of belts that were capable of accommodating waist sizes ranging from 24 inches to 38 inches. Also, I asked participants to wear cotton t-shirts to standardize the material interposed between the tactors and the skin.

I presented the tactile signal as a 250-Hz sinusoidal vibration presented with the same temporal profile and onset as the auditory stimuli (i.e., three pulse trains, each pulse 100-ms long, with a 50-ms silent interval between pulses one and two, and pulses two and three and a 500-ms silent interval after the third pulse). This signal frequency was chosen based on previous research showing that 250-Hz was the most localizable vibrotactile signal (Brill, 2015). I did not include a degraded cueing condition for the tactile stimuli. Additionally, due to the acoustic signature that the tactile belt created, (i.e., auditory sound generated from the tactors) I presented the auditory masker (diffuse white noise) during the tactile-only cue condition. This was not necessary for the auditory-tactile condition because the onset and offset of auditory and tactile stimuli were matched in this condition, and pilot study participants confirmed that the auditory stimuli and white noise were set at a level so as to mask any sound from the tactile belt. This ensured that the tactile-only cue did not involve any auditory spatial information.

Experimental Set-up

Participants completed all trials in a sound isolated room at AFRL/ WPAFB, sitting on a desk chair capable of 360 degrees of swivel rotation. There were three cords connected to the participant (i.e., the Oculus Rift cord, the Sennheiser headphone cord, and the EAI tactile belt cord). To ensure free-range of rotational (i.e., yaw) chair motion and vertical (i.e., pitch) head-motion for participants, I attached the Tactor Control Unit and tactile belt cord to the underside of the desk chair and instructed participants to keep the other two cords between their legs during chair rotation. Any excess cord was taped to the ground beside/ underneath the chair so that participants could rotate freely for 360 degrees during each trial. Participants were also instructed to return to their starting position by rotating the opposite the direction they rotated during the trial so as to avoid any cord entanglement. Participants were also instructed to indicate to the experimenter if cords became tangled during an experimental block by looking to the red virtual floor. The experimenter could monitor the participant's current status by viewing a computer screen that showed what the participant was seeing through the Oculus Rift. Therefore, a participant looking at an all-red virtual floor signaled to the experimenter that the participant needed assistance inside the sound isolated room.

Additionally, the Oculus Rift was equipped with Constellation, a room-scale head tracking system with a global position update rate of 60-Hz. The sensors responsible for maintaining head position information about the participant had a 1000-Hz sampling rate and a 500-Hz reporting rate. I integrated the visual, auditory, and tactile stimuli into the Constellation system to ensure that cues were updated as accurately as possible with participant body and head movement so that the cued spatial location of the visual target were well correlated with the actual location of the visual target. Participants were instructed to keep their head and torso fixed relative to one another during rotation and use their legs to turn the chair to explore the virtual environment so that the tactile belt would receive accurate spatial information from the participant's head position. Figure 5 shows a participant within the experimental set-up.

Figure 5. Front- and side-views of a participant in the experimental set-up. Participant is shown inside the sound isolated room with the Oculus Rift virtual reality headset and hand controllers, Sennheiser HD 280 professional headphones, and EAI C2 linear transducer tactors on the inside of an elastic belt with a Velcro fastener. The cables are shown taped to the floor under yellow and black caution tape.

Procedure

The research protocol and informed consent documents for this experiment were approved by the Wright State University and Wright-Patterson AFB Air Force Research Laboratory Institutional Review Boards. Detailed task instructions for the experimenter, including verbal prompts for participant task instructions, are included in Appendix A. Participants were fitted with the virtual reality headset, headphones, and tactile belt prior to each of the three experimental sessions. Participants were told to complete each trial as quickly and accurately as possible to ensure that neither speed nor accuracy was prioritized. To begin each trial, participants positioned their bodies and heads towards a center point at zero degrees elevation and zero degrees azimuth in the virtual world and focused on a fixation cross. Once the head tracker detected the correct position of the head, participants could hit a start button on the Oculus controller to begin the trial. When participants located the visual target, they pulled the trigger on the left or right Oculus controller to indicate the location of the gap on the target Landolt C. Response time was recorded from the onset of the trial until the trigger pull indicated target identification. Accuracy was recorded as correct or incorrect based on the congruency of the gap location and the button press.

Table 2 depicts the content of each of the three experimental sessions included in the experiment for one participant. The first session involved a two-part training. For Part 1 of training, participants completed one block of 10 trials per condition (i.e., no cues, auditory cues, degraded auditory cues, tactile cues, auditory with tactile cues, degraded auditory with tactile cues) with one visual target and 10 visual distractors. This small number of distractors (i.e., 10), allowed participants to build familiarity with the experimental set-up and task. During Part 2 of training, participants completed one block of 10 trials per condition with one visual target and 100 visual distractors. During training, the cues were presented in blocks according to the following order: no-cue, auditory cues, degraded auditory cues, tactile cues, audio-tactile cues, and degraded audio with tactile cues. I provided feedback to participants regarding their accuracy (i.e., “correct” or “incorrect”). I also prompted participants with the cue type that would be presented in the next block of trials. During the experimental sessions, the conditions were presented in random order across trials and across participants. After training in the first experimental session, participants completed two blocks of 15 trials per condition as part of the experimental trials. The second and third experimental sessions involved four blocks of 15 trials per condition each, for a total of 150 experimental trials per condition per participant and a total of 900 experimental trials per participant. Participants were given at least one hour of break time between sessions, and completed no more than two sessions per day.

Table 2

Experimental Session Content for One Participant

Session	Training Part 1	Training Part 2	Experimental Trials
Session 1	1 block x 10 trials per condition x 6 conditions = 60 trials	1 block x 10 trials per condition x 6 conditions = 60 trials	2 blocks x 15 trials per conditions x 6 conditions = 180 trials
Session 2			4 blocks x 15 trials per condition x 6 conditions = 360 trials
Session 3			4 blocks x 15 trials per condition x 6 conditions = 360 trials
		TOTAL	**900 experimental trials**

III. RESULTS

Excluded Trials

Three participants experienced multiple trials during which the tactile belt did not produce vibrations due to equipment malfunction. I ran all analyses both with and without the trials that were affected. When I ran the analyses without the affected trials, I excluded 225 trials for Participant 2, 360 trials for Participant 5, and 135 trials for Participant 10, one-third of which were tactile cued conditions, one-third of which were audio-tactile conditions, and one-third of which were degraded auditory with tactile conditions. Significant findings remained significant and non-significant findings remained non-significant when I included the affected trials. Therefore, the results presented exclude all trials that did not have a properly functioning tactile display.

Participant 5 reported that the auditory masker, which was intended to be diffuse noise, sounded as if its dB levels were changing with head movements. Upon further discussion, the participant confirmed that the level changes were occurring specifically during head movements that involved tilting one's head left or right so that the left or right ear, respectively, was facing the virtual floor. The study was designed for only rotational and elevation changes in head position. Therefore, the diffuse noise source was a virtual location directly below the participant's feet, and the level changes, although unexpected, are not surprising. No other participants reported these masker level changes. All of the trials in which Participant 5 reported hearing masker level changes were excluded, although the primary reason for the exclusion was due to tactile display malfunction. All raw data and analysis code are available upon request.

Accuracy Analysis

I expected that participants would show high accuracy overall. Participants were highly accurate, with an overall average of 99% correct responses. Mauchly's test indicated that the assumption of sphericity had been violated for auditory cueing conditions ($W = 0.521$, $p = 0.012$, $\varepsilon = 0.676$), and for the interaction of auditory and tactile cueing conditions, ($W = 0.263$, $p < .001$, $\varepsilon = 0.576$). Therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. A two-way analysis of variance showed that there was not a significant difference among participants in auditory conditions, $F(2,26) = 2.948$, $p = 0.070$, $\eta^2 = 0.044$, or in tactile conditions $F(1,13) = 1.179$, $p = 0.297$, $\eta^2 = 0.007$ (see Table 3). However, the interaction of auditory and tactile cueing conditions was significant, $F(2,26) = 3.879$, $p = 0.034$, $\eta^2 = 0.056$. The interaction between auditory and tactile cueing conditions indicates that the auditory cueing conditions had different effects on accuracy depending on the tactile cueing condition. As shown in Figure 6, when the auditory cues were absent, participants showed higher accuracy when tactile cues were present than when they were not.

Considering the two-part nature of the visual search task (i.e., localization followed by discrimination), there should be no significant effects regarding accuracy across conditions, because the same visual information is available for all conditions to use for the discrimination portion of the task. One possible explanation is that while the auditory and/or tactile cues were not assisting in the discrimination portion of the task, the benefit the cues provided during the localization portion ensured that participants completed the localization portion effectively, allowing for an accurate assessment of the visual target, rather than a guess, during the visual discrimination.

Table 3

Mean Accuracy Levels by Auditory and Tactile Cueing Conditions at the Group Level

	Auditory					
	Absent		Present		Degraded	
Tactile	*M (%)*	*SD*	*M (%)*	*SD*	*M (%)*	*SD*
Absent	0.967	0.123	0.993	0.060	0.988	0.083
Present	0.989	0.078	0.983	0.110	0.990	0.083

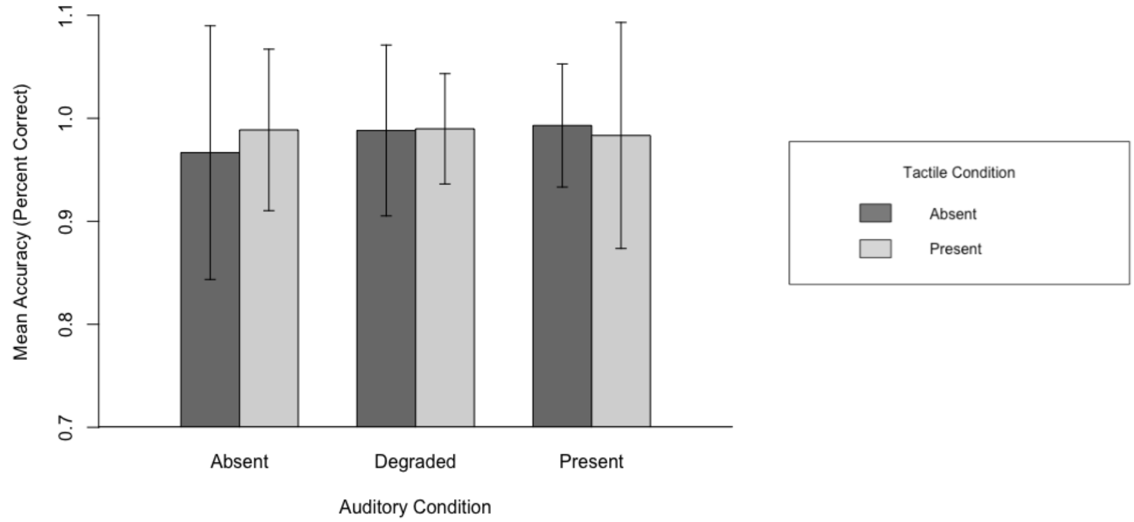

Figure 6. Mean accuracy levels plotted as a function of auditory cueing conditions. The parameter in the graph represents tactile cueing conditions. The error bars in the graph indicate averages of standard deviations across participants.

Mean Response Time Analysis

I hypothesized that there would be differences in mean RT across cueing conditions but that cueing of any form (i.e., auditory and/or tactile, present or degraded), would lead to significant reductions in visual search times over the no-cue condition. Table 4 below summarizes the group level mean RTs, SDs, and reductions of each cueing condition when compared to the no-cue condition.

A repeated measures analysis of variance was conducted on RTs for the auditory and tactile cueing conditions. Mauchly's test indicated that the assumption of sphericity had been violated for auditory cueing conditions ($W = 0.030$, $p < .001$, $\varepsilon = 0.510$), and for the interaction of auditory and tactile cueing conditions, ($W = 0.012$, $p < .001$, $\varepsilon = 0.504$).

Therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. There was a main effect for auditory cueing conditions on RT, $F(2,26) = 619.359$, $p <.001$, $\eta^2 = 0.876$, and a main effect for tactile cueing conditions on RT, $F(1,13) = 78.387$, $p < .001$, $\eta^2 = 0.540$. The interaction of RTs for auditory and tactile cueing conditions was also significant $F(2,26) = 66.519$, $p < .001$, $\eta^2 = 0.689$. The interaction between auditory and tactile cueing conditions indicates that the auditory cueing conditions had different effects on RT depending on the tactile cueing condition, as shown in Figure 7. These results supported my hypothesis that there would be differences in mean RT performance across cueing conditions.

Table 4

Summary of Mean Response Time Data

Cueing Condition	Mean Response Time (s)	Response Time Standard Deviation	Avg RT Reduction Compared to No-Cue Condition (s)
No-Cue	14.22	13.54	--
Auditory	2.29	1.27	11.93
Degraded Auditory	2.57	1.83	11.65
Tactile	4.87	4.47	9.35
Audio-Tactile	2.22	1.23	12.00
Degraded Auditory and Tactile	2.46	1.26	11.76

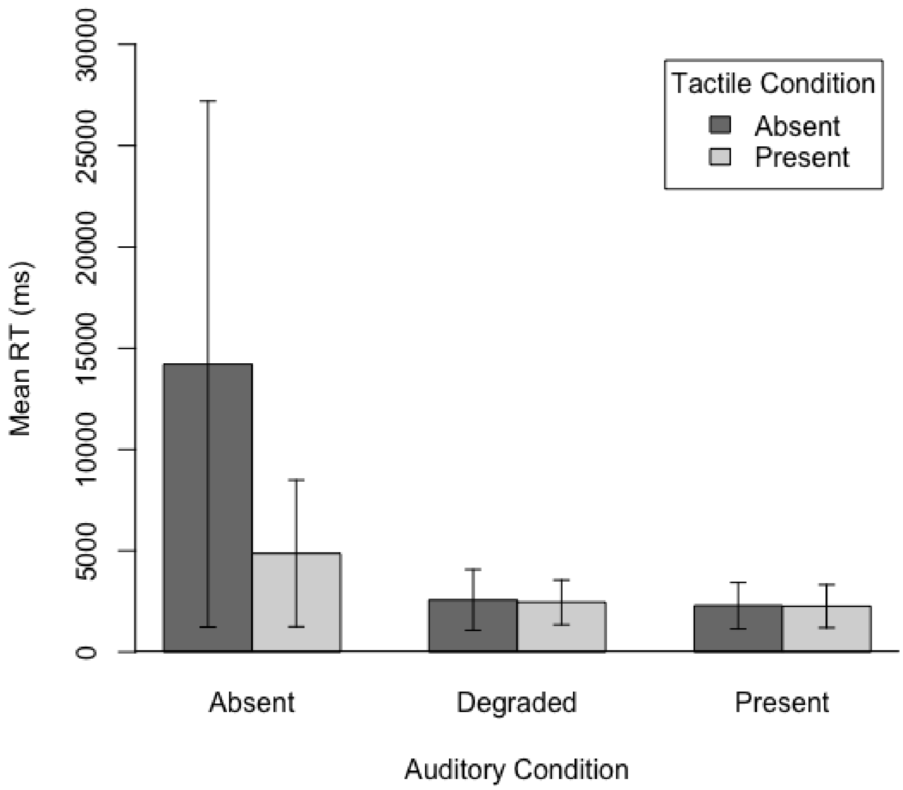

Figure 7. Mean RT levels plotted as a function of auditory cueing conditions. The parameter in the graph represents tactile cueing conditions. The error bars in the graph indicate averages of standard deviations across participants.

I completed Welch unequal variance two-sample t-tests to compare mean RTs for each condition involving (present or degraded) auditory and/or tactile cues to the no-cue condition. The test showed that mean RT was significantly lower for the auditory cue condition than the no-cue condition $t(1) = 37.82$, $p < .001$, $d = 1.24$. Mean RT was significantly lower for the degraded auditory cueing condition than the no cue condition, $t(1927) = 35.75$, $p < .001$, $d = 1.21$. There was a significant reduction in mean RT from the no-cue condition to the tactile cue condition, $t(2261) = 28.27$, $p < .001$, $d = 0.93$. The condition including both auditory and tactile cues yielded a significant reduction from the no-cue condition, $t(1890) = 37.93$, $p < .001$, $d = 1.24$. Finally, the condition including a combination of degraded auditory and tactile cues showed a significant reduction in mean RT from the no-cue condition, $t(1891) = 36.10$, $p < .001$, $d = 1.22$. These results supported my prediction that all cueing conditions would lead to significant reductions in mean visual search RT over the no-cue condition.

To extend the mean RT analysis beyond comparison to the no-cue mean, I compared mean RTs of cued conditions to other cued conditions. There was a significant difference between the auditory cueing condition and the degraded auditory cueing condition, $t(3311) = 5.40$, $p < .001$, $d = 0.18$. The mean RT reduction from the tactile condition to the auditory condition was also significant, $t(2157) = 23.90$, $p < .001$, $d = 0.78$. Mean RT was significantly reduced from the tactile condition to the degraded auditory condition, $t(2465) = 20.50$, $p < .001$, $d = 0.67$. These findings show that whereas cueing in general can enhance visual search performance in terms of mean RT, the type of cue affects the magnitude of the performance enhancement.

I predicted that the multimodal cueing condition (i.e., audio-tactile condition) would yield the fastest performance of all the conditions. However, the auditory cueing condition and the audio-tactile cueing condition showed nearly equivalent mean RT reductions over the no-cue condition, with the two conditions tying for the fastest RT performance (see Table 4). A Welch's unequal variance two-sample t-test confirms that the audio-tactile and auditory conditions showed no significant difference in mean RT,

$t(3714) = 0.77$, $p = 0.440$, $d = 0.03$. There was also no significant difference between the degraded auditory with tactile condition and the degraded audio condition $t(3294) = 2.30$, $p = 0.021$, $d = 0.075$. However, there was a significant RT reduction from the tactile condition to the audio-tactile condition, $t(2139) = 24.25$, $p < .001$, $d = 0.80$. There was also a significant RT reduction from the tactile condition to the degraded auditory and tactile condition, $t(2151) = 22.49$, $p < .001$, $d = 0.74$. This means that in terms of mean RTs, the tactile component of the audio-tactile cue provided no additional benefit over the auditory cue alone. Two types of additional analyses further examined this finding. First, a workload capacity analysis examined the cognitive processing speed for each particular cue condition. Second, examining trials that involve visual target onset outside of the participant's field of view and comparing RTs from tactile-cue conditions to non-tactile cue conditions allowed for a greater understanding of the benefits (or lack of benefits) provided by the tactile display. Both analyses are explored below.

Workload Capacity Analysis

I performed workload capacity analyses in order to examine the entire range of RTs, rather than only the mean RTs, and also to examine the magnitude of improvement from single-modality and multimodality cued conditions (i.e., improvement from auditory condition or tactile condition to audio-tactile condition) and its cognitive processing implications. First, I examined each cueing condition using the no-cue condition as the baseline for performance for each individual using the "ST-ST" capacity coefficient workload capacity analysis, which yields capacity Cz-scores, to examine the full range of RTs. Table 5 below shows the capacity Cz-scores for each individual participant across cued conditions, with the no-cue condition as the baseline, with a single asterisk indicating performance equal to the baseline and a double asterisk indicating worse performance than the baseline.

Table 5

Capacity "ST-ST" Cz-scores for each Participant Across Cued Conditions

Participant Number	Auditory Condition	Degraded Auditory Condition	Tactile Condition	Audio-Tactile Condition	Degraded Auditory w/ Tactile Condition
1	7.35	6.91	5.41	7.52	6.95
2	2.34	2.95	2.18	2.82	4.03
3	8.74	7.78	4.22	8.72	8.32
4	7.88	7.85	4.23	8.15	7.31
5	3.60	1.91	2.06	2.87	1.71
6	9.89	9.60	6.81	9.83	9.68
7	6.47	5.67	3.79	6.56	5.91
8	7.10	6.50	3.32	7.25	6.56
9	6.32	5.47	-1.66†	6.28	6.16
10	6.19	5.63	3.05	6.16	5.78
11	8.11	7.28	4.61	8.19	7.58
12	8.61	7.83	5.52	8.56	7.91
13	9.46	8.92	4.71	9.22	8.81
14	9.30	8.56	6.78	8.94	8.38

Note: A single dagger(†) indicates worse performance than the baseline no-cue condition.

Figure 8 shows that the capacity coefficients of most individual participants were well above the no-cue baseline, showing super capacity processing for all cueing conditions. Functions higher than the baseline show better performance relative to the no-cue condition. Functions lower than the baseline show decreased performance relative to the no-cue condition. All participants show performance better than the baseline for all cue conditions except Participant 9, who shows performance worse than the baseline in the tactile condition. The figure shows that the mean RT reductions were robust to the entire range of RTs.

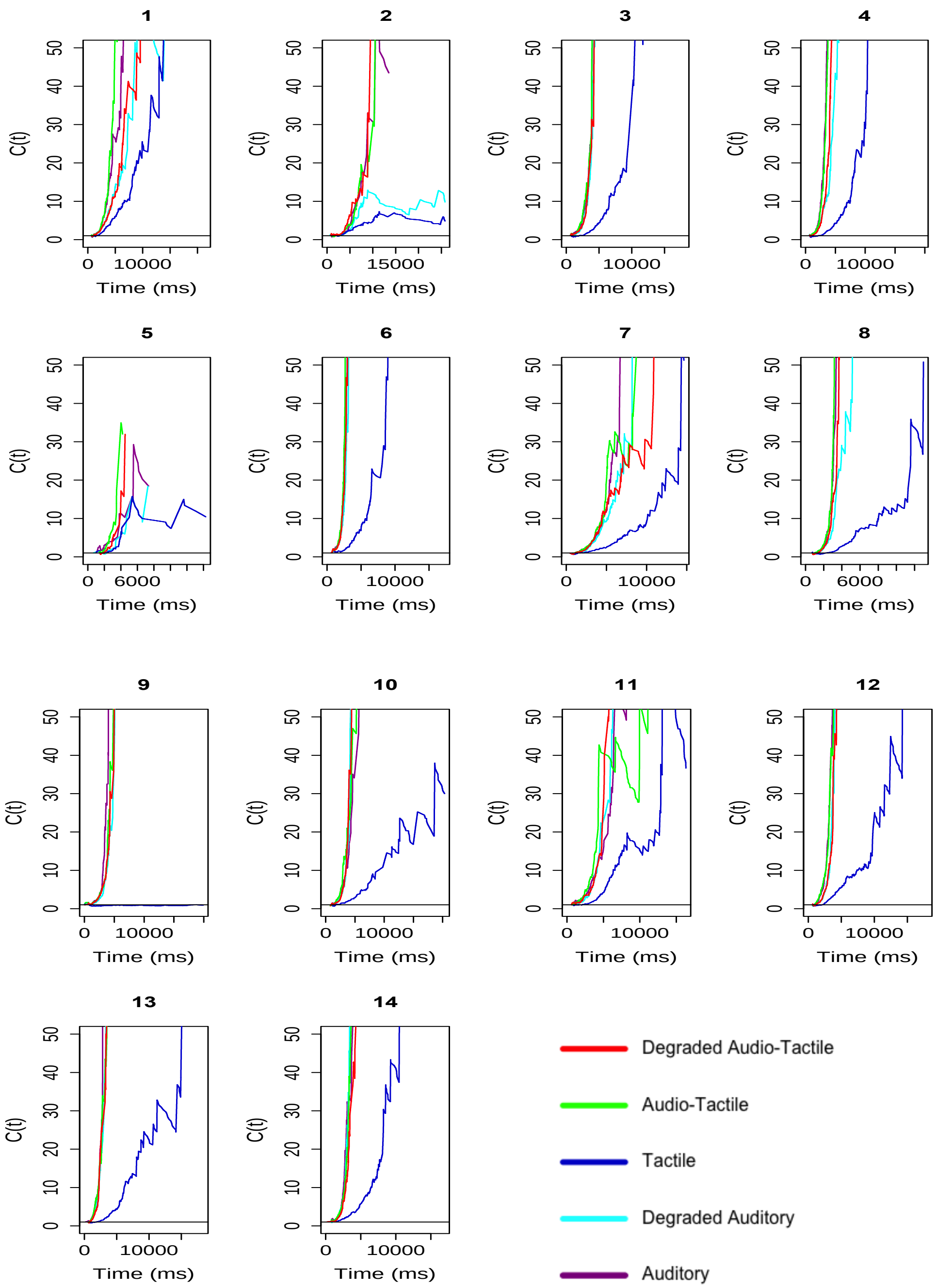

Figure 8. Capacity "ST-ST" coefficients for each participant across conditions. The title of each graph indicates participant number. The y-axis represents multiples of betterness compared to the baseline model. The black line at y = 1 indicates the baseline no-cued condition for comparison.

I examined the capacity "OR" coefficients, which yielded capacity functions for all participants with the UCIP model as the baseline for comparing the single-modality conditions (i.e., auditory and tactile) to the multimodal cueing conditions (i.e., the audio-tactile). Figure 9 shows that the group mean compared to the UCIP model baseline at *C(t)* = 1 is unlimited capacity. Most participants, as well as the group average, were below the UCIP model baseline, showing limited capacity.

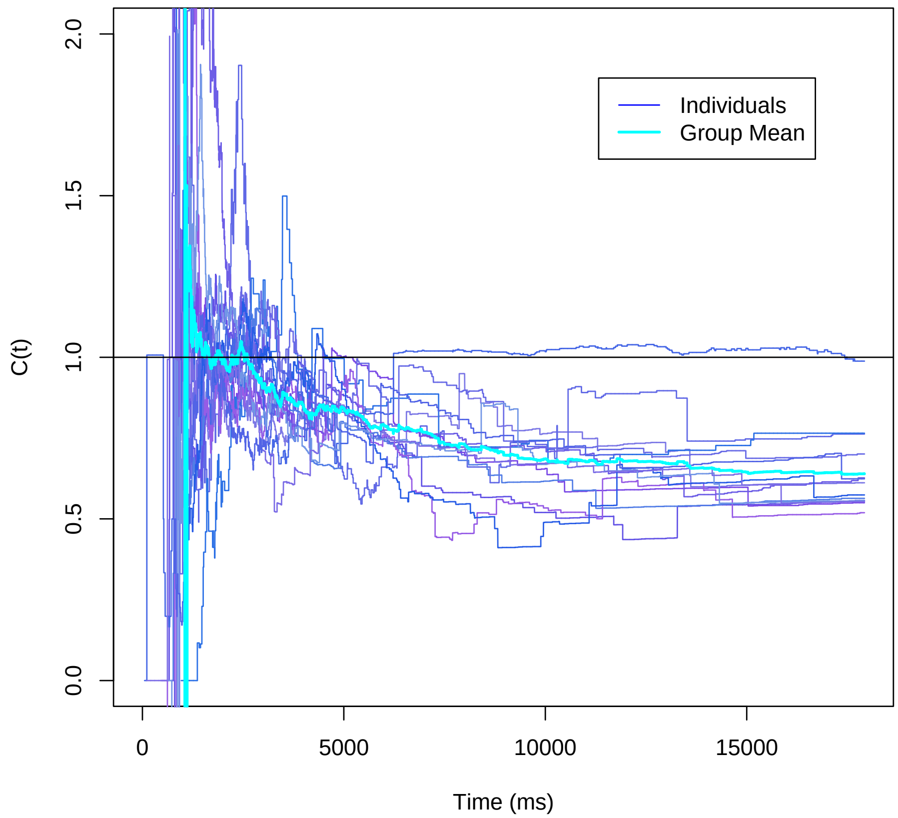

Figure 9. Capacity "OR" coefficients for each participant and for the group mean. The x-axis represents time in milliseconds. The y-axis represents multiples of better-ness compared to the UCIP baseline model of statistical facilitation, indicated by the black line at y = 1.

Analysis of Target Onset Locations Behind Participant

As a follow-up analysis to the finding that the tactile display did not provide an additional RT benefit when auditory cues were present, I examined trials for which the visual target onset location was behind the participant (i.e., target onset location greater than or equal to +/- 90 degrees azimuth from participant starting head position). I used a repeated measures analysis of variance to compare RTs for tactile cue form (i.e., absent or present) and auditory cue form (i.e., absent, present, or degraded) when the visual target onset was behind the participant. Mauchly's test indicated a violation of the assumption of sphericity for the auditory cueing conditions ($W = 0.041$, $p < .001$, $\varepsilon = 0.510$), and for the interaction of auditory and tactile cueing conditions, ($W = 0.022$, $p < .001$, $\varepsilon = 0.510$). Therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. There was a main effect of auditory cue form on RT, $F(2, 26) = 412.13$, $p < .001$, $\eta^2 = 0.842$, and a main effect of tactile cue form $F(1, 13) = 87.81$, $p < .001$, $\eta^2 = 0.534$. The interaction between auditory cue form and tactile cue form was also significant $F(1, 13) = 76.43$, $p < .001$, $\eta^2 = 0.69$. The significant interaction indicates that the tactile cue form had different effects on RT depending on the trial onset location, as shown in Table 5 and Figure 10. As in the mean RT analysis for all trials, the no-cue condition shows a large standard deviation, likely due to the variation in target onset locations (i.e., close to participant's field of view at trial onset or far from participant field of view at trial onset).

These results are similar to the initial findings that the tactile display shows the biggest RT benefit when no other cues were present. However, the magnitude of this benefit is greater only when examining trials for which target onset location was behind the participant. This supports the literature that visual and tactile modalities have complementary benefits and that when visual targets are outside the participant's field of view, the tactile modality becomes more important if no other cues are provided. However, the finding is insufficient to determine that the tactile cueing modality should be provided even if auditory cues are present.

Table 6

Means and Standard Deviations of RTs for Target Onset Location Behind Participants

	Auditory					
	Absent		Present		Degraded	
Tactile	*M (s)*	*SD*	*M (s)*	*SD*	*M (s)*	*SD*
Absent	16.79	12.41	2.73	0.98	3.03	1.32
Present	5.62	3.97	2.71	1.03	2.90	1.03

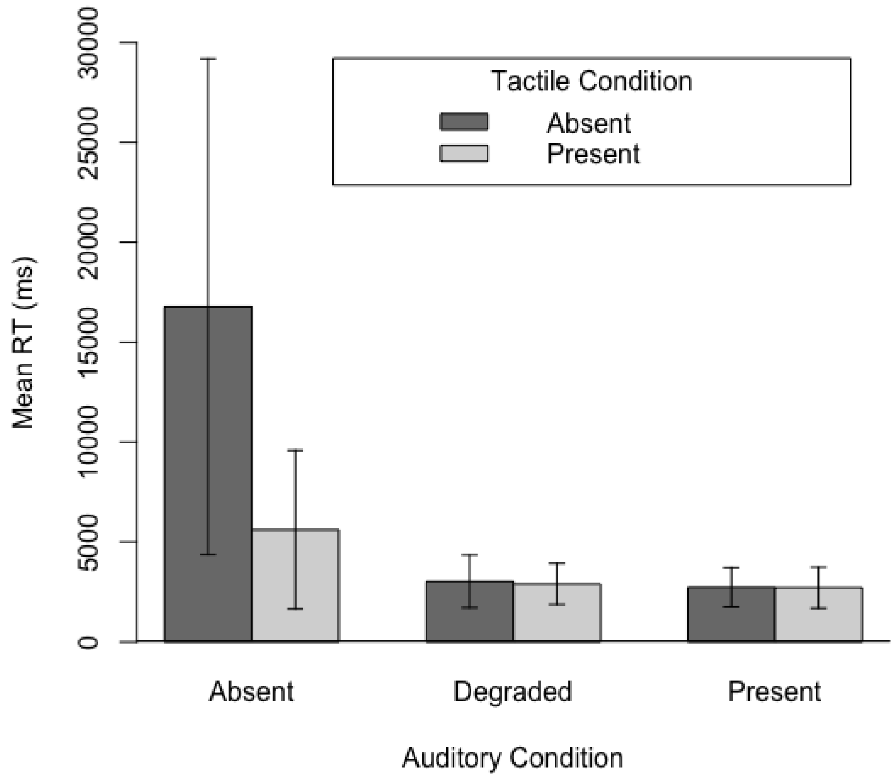

Figure 10. Mean RT levels for target onset location behind participant plotted as a function of auditory cueing conditions. The parameter in the graph represents tactile cueing conditions. The error bars in the graph indicate averages of standard deviations across participants.

IV. DISCUSSION

The purpose of my research was to evaluate dynamic visual search performance in a 3D virtual environment when spatialized auditory and/or vibrotactile cues are provided and to examine the cognitive processes involved. Participants showed high accuracy overall, but the no-cue condition showed significantly lower accuracy than the cued conditions. Both the mean RT analysis and the workload capacity analysis showed clear evidence of improved speed in a dynamic visual search task when non-visual cues were provided, demonstrating the robustness of spatialized auditory, tactile, and audio-tactile cueing to three-dimensional environments. This is an important finding because it shows the utility of spatialized auditory and tactile cueing beyond simply reducing the initial search space, showing that spatialized auditory and tactile cueing can be used to provide continuously updated spatial information when targets are moving.

The significant difference between accuracy in the no-cue condition to the cued conditions is interesting because the same visual information (i.e., the information used to make the visual discrimination) was present for all conditions. It is possible that while the cues (i.e., auditory and/or tactile) were not assisting in the discrimination itself, the benefit the cues provided during the localization portion ensured that participants completed the localization portion on most trials for the visual discrimination portion of the task. The fact that the no-cue condition shows the lowest amount of accuracy might show support for this claim. The surprising significant interaction between auditory and tactile accuracy levels might mean that during the no-cue condition (when both auditory and tactile cues were absent), sometimes participants were not completing the localization portion of the visual search task, and therefore cannot accurately complete the discrimination portion of the task (i.e., the participants simply guessed the gap location), resulting in a lower level of accuracy for the no-cue condition than the cued conditions at the group level. This might be because no-cue conditions generally required more search time than other conditions and therefore the time the participant expected to spend on the trial might have been exceeded. The participants might have "bailed-out" of the trial due to boredom or frustration at not being able to locate the visual target within their expected time frame. Further examination of the cognitive processing involved in the localization portion of the visual search task is examined in the context of RTs in the mean RT and workload capacity analyses below.

The significant differences between the forms of cueing conditions (i.e., present, absent, or degraded) showed that the quality of the cue can affect performance differentially, with the degraded cueing conditions yielding worse performance than the not degraded (i.e., present) conditions. This finding is useful in two ways. First, in multimodal research settings, the level of cue degradation should depend on the level of cross-modal matching desired. The results of this study showed that a spatialized auditory cue that is degraded with a diffuse white noise auditory masker is not degraded enough to match the lower level of effectiveness of a spatialized tactile cue. This might indicate that participants were relying on the auditory cue, regardless of its degradation, implying one of two things: either the auditory cue was not degraded enough, or the tactile cue was not effective enough. If the former, future research should provide auditory cues that have a lower signal-to-noise ratio, or degrade the auditory display in another way, rather than using an auditory masker. One way to do this is to provide inaccurate spatial information, as shown by Haggit (2014). Another is to delay spatial information presentation. If the latter, future research should use a tactile display that provides higher spatial resolution.

The second way that understanding the implications of cue quality is important is that in operational settings, display degradation may occur unintentionally. The results of this study show that masking a spatial auditory signal with auditory noise does affect the utility of auditory cueing in a visual search task. Applied to an operational setting such as aviation, this might mean that a noisy airplane cockpit might mask some auditory cues. Although this experiment did not include a degraded tactile display condition, it is reasonable to assume that tactile masking might affect the utility of a directional tactile cue in a visual search task in a way similar to auditory noise, particularly if the tactile cues and tactile noise provide vibration in the same frequency range (Brill, 2015). For example, turbulent flight conditions might mask some tactile cues, reducing their effectiveness for providing spatial information. Because the results of this study showed that the degraded displays can affect the utility of auditory cues in visual search tasks, future research should examine the magnitude of these effects under conditions that commonly occur in operational settings. Hopefully, this research will inspire the creation of improved displays that are effective even under harsh conditions so that operational effectiveness will not be in jeopardy.

Additionally, the standard deviation for the no-cue condition (i.e., auditory absent and tactile absent), is particularly large. This is likely due to the effect of visual target onset location. When there was no auditory or tactile cue to assist participants in localization, participants had to rely solely on visual information to locate the target. Intuitively, the targets that were further away from the participant's initial field of view took longer to detect than those that were closer, because there was more space to search. This means that in the no-cue condition, there were some trials with a relatively long RT and some with a relatively short RTs. Therefore, the range of response times (and consequently the standard deviation) was larger in the no-cue condition than the other conditions. Taking the visual target onset and continuously updated locations into account would allow for a greater understanding of the effects of auditory and tactile cueing on visual search performance. I plan to complete this analysis in a future project, using the streaming head-position and target data that I collected during this experiment.

The capacity "ST-ST" analysis showed that these results were robust across the entire range of RTs (i.e., that providing cues made visual search more efficient across the range of RTs). The capacity "OR" analysis showed that the addition of the tactile cue to the auditory cue did not improve processing speed for the individual cues as predicted by the UCIP model but instead slowed the processing speed for the individual information sources. These results provided evidence that whereas spatialized multimodal (i.e., audio-tactile) displays can enhance dynamic visual search, spatialized auditory displays can provide just as much or more cognitive processing benefit in these experimental conditions. To make this psychologically meaningful, the results of the capacity "OR" analysis show that the addition of multiple modalities (i.e., auditory and tactile) to the visual modality involved in the visual search task decreased the processing speed for each individual modality, indicating statistical debilitation.

Explanations for Tactile Display Performance Outcomes

Given that the tactile display was not as effective as the auditory display, the nature of tactile cueing merits further exploration. There are two major categories of explanation to explore. First, the tactile display was not as effective as expected because participants did not process the cue as expected, due to lack of familiarity or cognitive architecture. Second, the tactile display was not as effective as expected because the display itself was not as effective as intended, due to cross-modal spatial information presentation mismatch with auditory or visual displays.

Participant familiarity with auditory displays. One possibility for the finding that the tactile display did not provide as much RT benefit as expected was that participants were more familiar with the auditory display than the tactile display. In general, most people have used headphones, and might have even experienced spatialized sounds through headphones. In fact, ten out of 14 of my participants came from a subject panel for an auditory research branch of the Air Force Research Laboratory. This means that most of my participants had experienced studies involving spatialized auditory sounds. However, none of the participants had experienced the sensation of a tactile belt. In fact, several participants commented that the vibrations on the torso were unique or interesting. Despite this cross-modal familiarity difference, the explanation for the lack of substantial tactile findings is more nuanced than what can be explained through participant preferences.

Parallel cognitive processing. The finding that the tactile display did not add any additional RT benefit to the auditory display might mean that processing of visual information inherent in the task and information from the auditory and/or tactile cues was not completed in parallel, as the UCIP model suggests. If cues are processed serially (i.e., one after another), rather than in parallel, then the addition of more cue types would never add an RT benefit, rendering multimodal displays unnecessary. As this seems an unlikely explanation since several studies have shown the benefit of presenting multimodal displays, it is possible that such a finding could be true for this particular type of visual search (i.e., dynamic and 3D). However, there were participants who showed unlimited capacity processing when presented with the multimodal cues. Those participants might be following the UCIP model, meaning that they were processing the cues in parallel, which is why there is no performance benefit for the tactile cue.

Cross-modal mismatch for spatial information presentation. Another potential reason for this lack of substantial tactile display benefit in the multimodal conditions is that the cues were not well matched in terms of spatial information provided. Although the auditory and tactile sensory systems have complementary strengths and weaknesses, the auditory cues were more effective at providing spatial information than the tactile cues in this experiment. This is not completely surprising considering two major differences in the auditory and tactile displays. First, the shape of the head maps onto a sphere better than that of a torso simply due to human anatomy.

Second, as a result of the previous point, the spatialized auditory cues presented both azimuth and elevation cues whereas the tactile cues were only capable of presenting azimuth cues every 45 degrees, which only changed when a dynamic cue traveled far enough to trigger the adjacent 45-degree cue. The tactile display employed in this research was originally developed for providing short-duration static azimuth cues, and as such, it could not deliver spatially or temporally accurate analogs to the auditory and visual cues. Lastly, there were several instances of tactile display malfunction. Whereas participant reports and/or software program reports alerted these malfunctions, it is possible that there were malfunctions of which I was not aware. However, I do not expect that the results would be changed drastically, as the inclusion/exclusion of the known tactile malfunction trials did not significantly affect the experiment results.

One possible follow-up study could use a virtual space that maps more efficiently onto the human torso. My experiment did attempt to address this issue by including floor and ceiling cut-offs for the visual stimuli trajectories because a target directly above or below the head would be difficult to display using a tactile belt that only provides azimuth cues. Another line of research could include a full-torso tactile display instead of simply a belt to address the lack of tactile elevation cues although the problem of the head and torso mapping onto a sphere differently would still be present. Another possible study could use the same auditory and tactile displays but examine dynamic visual search within a fixed elevation while varying azimuth locations. Although such an experiment would have less ecological validity to operational settings that involve omnidirectional visual search, the display designs would ensure that the amount of spatial information is well matched across modalities.

Another possible explanation is that there was enough spatial visual information within the participants' field of view that participants did not need to rely on the tactile cues for spatial information. Previous research has shown that tactile cues are most important when there is a threat degraded auditory cues through front-back reversals (i.e., when the visual target's onset location is in front of or behind the participant and is not within the visual field of the participant at trial onset; Brill et al., 2014). Therefore, I expected to see stronger RT benefits when both auditory and tactile cues were present for the trials in which the target onset location was behind the participant. However, the tactile cue did not provide as much of a benefit as the auditory and tactile modality complements would have predicted. This might be because any front-back confusion that came from the auditory display was disambiguated by visual information when target onset location was within the participants' field of view. A follow-up study involving a non-visual task, but instead a simple sound and/or vibration localization task would eliminate the possibility of using visual cues to disambiguate auditory (or tactile) confusion. Such research could provide a more extensive understanding of the complementary benefits of auditory and tactile modalities. Another possibility is that the auditory display performed better relative to the tactile display than predicted because the dynamic nature of the task meant that head motion and dynamic auditory cues disambiguated any spatial ambiguity quickly. This would imply that the dynamic or static nature of a visual search task is of utmost importance and should be considered in the creation of creation of multisensory displays.

The limitations of the tactile display in this experiment point to a need for more tactile cueing research. It is clear from the present results that there is no additional benefit of including a tactile display during a visual search task that does not involve visual targets outside of the field of view. However, when the task does involve such targets, the benefit of tactile cues requires further exploration. The practical implication of these limitations is that when considering the creation of a multimodal display for a task in which there is potential for dynamic targets to be outside the field of view of the operator, an auditory display cue should be provided if possible. If the visual targets will be behind the operator, tactile displays should be considered, depending on the constraints of the environment. However, if there is no reason to expect that targets will be located directly behind an operator, spatialized auditory cues are likely sufficient.

Conclusion

The primary goal of my research was to evaluate auditory and tactile enhancement of visual search. To do this, I expanded the current visual search research in two ways. First, I provided spatialized auditory, tactile, and audio-tactile cues for a visual search task in a 3D, dynamic environment and found that RT benefits of both auditory and tactile cueing over a non-cued visual search are robust to a 3D, dynamic environment. Second, I examined the cognitive processes involved in visual search using novel statistical techniques and found that including both auditory and tactile displays in the multimodal condition provided no additional benefit over the auditory condition in terms of mean RT or cognitive processing speed. In fact, I found that in the multimodal condition, processing of auditory and tactile sources individually was actually slowed as compared to the single-modality conditions. This finding means that, at least in audio-tactile dynamic visual search, auditory and tactile displays might be processed serially, as opposed to the UCIP baseline model prediction of parallel processing. Serial processing would mean that multimodal cueing would not be beneficial in an operational setting, in which processing speed is critical for effective performance.

This deeper understanding of the underlying cognitive architecture involved in processing of dynamic visual search tasks allows us to develop and implement multimodal information displays, such as spatialized auditory displays for visual tasks, according to specific cognitive processing and visual search needs. Consequently, future multimodal research and multimodal displays should consider that cognitive processing speed of visual search tasks in the context of auditory, tactile, or audio-tactile displays can be better than visual search tasks alone. In operational settings, the visual search enhancements provided by the addition of appropriate multimodal cues as demonstrated by this experiment will be of critical importance for operational effectiveness and might even save lives.

V. REFERENCES

Amor, T. A., Reis, S. D. S., Campos, D., Herrmann, H. J., & Andrade Jr., J. S. (2016). Persistence in eye movement during visual search. *Scientific Reports, 6.*

Bolia, R. S., D'Angelo, W. R., & McKinley, R. L. (1999). Aurally aided visual search in three-dimensional space. *Human Factors, 41*, 664-669.

Brill, J. C. (2015). Guidelines for vibrotactile display design: Front-end, intensive, and spatial considerations. In R. Hoffman, P. A. Hancock, R. Parasuraman, J. L. Szalma, & M. Scerbo's (Eds.), *The Cambridge Handbook of Applied Perception Research,* Cambridge University Press, New York.

Brill, J. C., Lawson, B. D., & Rupert, A. H. (2014). Tactile Situation Awareness System (TSAS) as a compensatory aid for sensory loss. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 58, No. 1, pp. 1028-1032). SAGE Publications.

Brill, J. C., Lawson, B., Rupert, A. H., & Gagliano, I. (2015). Audiotactile displays for improving situation awareness and mitigating spatial disorientation. *Proceedings of the American Helicopter Association,* Virginia Beach, VA.

Brill, J. C., & Scerra, V.E. (2014). Effectiveness of vibrotactile and spatial audio directional cues for USAF Pararescue Jumpers (PJs). *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics*, Krakow, Poland 19-23 July. Edited by T. Ahram, W. Karwowski and T. Marek.

Brungart, D. S. & Simpson, B.D. (2008). Effect of temporal fine structure on the localization of broadband sounds: Potential implications for the design of spatial audio displays. In *Proceedings of the 14th International Conference on Auditory Displays*.

Cholewiak, R. W., Brill, J. C., & Schwab, A., (2004). Vibrotactile localization on the abdomen: Effects of place and space. *Perception & Psychophysics, 66*(6), 970-987.

Endsley, M.R. (2000). Theoretical underpinnings of situation awareness. In M.R. Endsley & D.J. Garland (Eds.), Situation awareness analysis and measurement (pp. 1–21). Mahwah, NJ: Erlbaum.

Flanagan, P., McAnally, K. I., Martin, R. L., Meehan, J. W., & Oldfield, S. R. (1998). Aurally and visually guided visual search in a virtual environment. *Human Factors, 40,* 461-468.

Fox, E. L., Glavan, J. J., Houpt, J. W. (2014). Evaluation of cognitive processing in redundant audio−visual signals. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *36*(36).

Good, M. D. & Gilkey, R. H. (1996). Sound localization in noise: The effect of signal-to-noise ratio. *The Journal of the Acoustical Society of America, 99(2)*.

Haggit, J. (2014). *Cued visual search and multisensory enhancement.* Unpublished master's thesis. Wright State University, Dayton, OH.

Hancock, P. A., Lawson, B. D., Cholewiak, R., Elliot, L. R., van Erp, J. B., Mortimer, B. J. P., Rupert, A., & Redden, E. S. (2013). Tactile cuing to augment multisensory human-machine interaction. *Ergonomics in Design, 23(2)*, 2-9.

Houpt, J. W., Blaha, L. M., McIntire, J. P., Havig, P.R., & Townsend, J. T. (2014). Systems factorial technology with R. *Behavior Research Methods, 46*, 307-330.

Kingsley, H. L. (1932). An experimental study of 'search.' *American Journal of Psychology, 44*, 314-318.

Mateo, J. C., Simpson, B. D., Gilkey, R. H., Iyer, N., & Brungart, D. S. (2012). Spatial multisensory cueing to support visual target-acquisition performance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 56, No. 1, pp. 1302-1316). SAGE Publications.

McIntire, J.P., Havig, P. R., Watamaniuk, S. N. J., & Gilkey, R. H. (2010). Visual search performance with 3-D auditory cues: Effects of motion, target location and practice. *Human Factors, 52(1)*, 41-53.

Perrot, D. R., Cisneros, J., McKinley, R. L., & D'Angelo, W.R. (1996). Aurally aided visual search under virtual and free-field listening conditions. *Human Factors, 38,* 702-715.

Rupert, A.H. (2000). Tactile Situation awareness system: Proprioceptive prostheses for sensory deficiencies. *Aviation, Space, and Environmental Medicine, 71(6),* 571-578.

Santangelo, V. & Spence, C. (2007). Multisensory cues capture spatial attention regardless of perceptual load. *Journal of Experimental Psychology: Human Perception and Performance, 33(6)*, 1311 -1321.

Sherrick, C. E., Jr. & Cholewiak, R. W. (1986). Cutaneous sensitivity. In K. Boff, L. Kaufman & J. L. Thomas (Eds.), *Handbook of perception and human performance*, pp. 12-58. New York: Wiley.

Simpson, B. D., Brungart,, D. S., Dallman, R. C., Joffrion, J., Presnar, M. D., & Gilkey, R. H. (2005, September). Spatial audio as a navigation aid and attitude indicator. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 49, No. 17, pp. 1602 -1606). SAGE Publications.

Stein, B.E., Wallace, M.T., & Stanford, T.R. (2001). Brain mechanisms for synthesizing information from different sensory modalities. In E.B. Goldstein (Ed.), *Blackwell handbook of perception* (pp. 387-431). San Diego, CA: Academic Press.

Stokes, A. & Wickens, C. (1989). Aviation displays. In E. Wiener & D. Nagel (Eds.), *Human factors in aviation* (pp. 387 -431). San Diego, CA: Academic Press.

Townsend, J. T. & Wenger, M. J. (2004). A theory of interactive parallel processing: New capacity measures and predictions for a response time inequality series. *Psychological Review, 111(4)*, 1003-1035.

Van Erp, J. B. (2007). *Tactile displays for navigation and orientation: Perception and behavior*. Leiden, The Netherlands: Mostert & Van Onderen.

Van Erp, J. B., Groen, E. L., Bos, J. E., & van Veen, H. A. (2006). A tactile cockpit instrument supports the control of self-motion during spatial disorientation. *Human Factors, 48,* 219-228.

Wandell, B. A. (1995). *Foundations of Vision.* Sunderland, MA: Sinauer Associates, Inc.

Wolfe, J. (1998). What can 1 million trials tell us about visual search? *Psychological Science, 9(1),* 33-39.

APPENDIX A

Task Instructions for Experimenter (Including Prompts with Participant Task Instructions)

1 day before task (send email or give verbal instructions to participant):

- Wear contacts if necessary and/or possible but if not, then glasses are acceptable.
- Wear a shirt with cotton tshirt-like material (short or long sleeves is acceptable).
- Provide informed consent and obtain participant signature.
-

Before instructions (for experimenter):

- Set z file for Train "1".
- Measure waist with tape measure.
- Attach smaller or larger belt accordingly (24" – 31" gets smaller belt, 31" – 38" gets larger belt).

NOTE: Prompts to be read aloud by experimenter are in quotations(" "). Actions to be completed by experimenter are indented and in square brackets ([ ]).

Task Instructions (for experimenter to read to participant):

"You will be searching for a 'C' among distractor 'O's. Some of the Cs will have an opening facing to the right, and others will have an opening facing to the left. Your goal is to locate the C and determine the direction of the gap opening (left or right). To help you achieve this, sometimes you will be provided with cues. There will be some conditions with spatialized sound or auditory cues, presented through head phones. This means that if you hear a sound cue to your right, the target C is to your right. Some conditions will include spatialized vibration or tactile cues, presented through a vibrating belt that I will show you momentarily. This means that if you feel a vibration on your right side, the target is to your right. Some conditions will include both auditory and tactile cues. Some conditions will also include auditory white noise, which sounds like static. There will also be conditions that include no cues, so you will use only the visual environment to respond. I will further explain the cues once you are positioned in the experimental set-up."

[Give training or experiment instructions appropriately.]

Part 1 Training Instructions:

"Please have a seat on the chair and do not move the chair on the floor, but try spinning around. During the experiment, you will need to spin in both directions. Keep your head fixed with your torso (rotate with your whole body, don't just turn your head). Now, please pick up the tactile belt. You will wear it about 1 inch above your navel, with the Velcro centered above your belly button. In a moment, I will demonstrate a test vibration. During the experiment, the location of the tactile cue indicates the spatial location of the target 'C'. For example, if you feel a vibration on your right side, that means the target 'C' is to your right and you should spin that way to locate it."

[START EXPERIMENT WITH DOOR OPEN and z-file with following conditions:

- Participant #: [insert participant number]
- Trials: 10
- Train: "1"

- Once you have run the Unity file, say "You should feel the test vibration in a moment."
- Go back to sound booth to help participant with remaining equipment.]

"Next I will explain the hand controllers. You will have one in each hand. To start each trial you will look forward until you see the crosses align (you'll see this in a minute). Then you will start the trial by using your right thumb to hit the 'A' button on the controller here."

[Point to 'A' button]

"Once the trial starts, you will see the visual stimuli moving around and you should look for the target 'C' and notice the direction of its gap. When you have located the target 'C', if the gap is facing to the right like this…" [use hand to show a 'C' with gap opening facing to participant's right] "… then you would pull the right-hand inner trigger like this." [Demonstrate pulling right inner trigger.] if the gap is facing to the left like this…" [use hand to show a 'C' with gap opening facing to participant's left] "… then you would pull the left-hand inner trigger like this."

[Demonstrate pulling left inner trigger, confirm that participant understands, and then set controllers down to explain remaining equipment].

"Before I explain the rest of the equipment, note that there are several cords in this set-up. Keep the cords between your legs in the front and before each trial, come back to center the same way you left.

"Next I will help you put on the Oculus Rift."

[Before handing Oculus to participant, show how to adjust on sides and top, then help as necessary].

“You should now see a gray background with ‘O’s moving around and white text on the screen. Look down now to see the red floor.” The target ‘C’ and distractor ‘O’s will never be lower than the edge of the floor, and you will not have any directly above you on what we call the ‘ceiling’ either. Look up now to see how high they are moving as an example.

“If you see a blue grid pattern like this…

[pick up one controller and hold it close to the wall]

…that is the boundaries of the physical environment (the wall), so keep your arms and legs inside the grid or you may hit the wall.

“If at any point you feel you need to leave the experiment or you need help with the equipment (cords tangled, etc.), look straight down so that your field of view is entirely covered by the red floor and I will come help you. Try to finish your current trial if possible.

“Before I help you with the headphones, I will explain Part 1 of training. For Part 1 of training, you will have 1 target (the ‘C’) and 10 distractors (the ‘O’s) for each trial. You will complete 10 trials each for six cueing conditions, presented in this order:

- No cues
- Auditory cues
- Auditory + masker (noise) cues
- Tactile + masker (noise) cues
- Tactile + auditory cues
- Tactile + auditory + masker (noise) cues

“You will also receive on-screen feedback about your responses (i.e., "Correct" for correct responses or "Incorrect" for incorrect responses) for the first part of training. Once you have completed Part 1 of training, stay where you are and I will give you more instructions for Part 2 of training. Remember that you are searching for a ‘C’ among the ‘O’s and you will identify the direction that the gap opening on the ‘C’ is facing.

“Next I will explain the auditory cues and then hand you the headphones followed by the controllers. Then I close the door and you can begin the first trial whenever you are ready. Auditory cues indicate the location of the target ‘C’ in the same way that the tactile cues do. That means if you hear a sound from the right, the target is to your right, so you should swivel that way.”

[Once headphones are on (cord on left ear), place controllers in participant’s hands, visually confirm that all equipment is fitted correctly, then close the door. If participant does not begin trial within ~20 seconds open the door and confirm they understand that they can begin. Once they begin, monitor the computer screen (which shows their field of view) to ensure they are completing the trials properly and watch for an entirely red screen. If you see an entirely red screen, immediately open the sound booth door and ascertain source of problem from the participant.]

Part 2 Training Instructions:

[After participant has completed Part 1 of training (6-7 minutes), open the door. If participant does not remove headphones, tap them on the shoulder to prompt them to do so.]

"You have completed Part 1 of training. Now we will move into Part 2 of training. For this part of training, you will see 10 trials of each of the 6 conditions in the same order as Part 1 of training. However, for Part 2, there will be 1 target and 100 distractors present. You will not receive accuracy feedback for this part of training. Remember there is a target present in every trial. I'll close the door and then start Part 2 of training. You will feel another test vibration and once you see the experiment, you may begin."

[Set z-file for Train "2" and then begin training Part 2 and monitor the screen throughout.]

Experiment Instructions:

If Session 1 after training Part 2:

"You have completed the training portion of this experiment. Do you have any questions or need a break? If not, we will move into experimental trials. Each trial will be similar to Part II of training in that there will be 1 target 'C' and 100 distractor 'O's. Remember there is a target present in every trial. However, the order of conditions will be random throughout the experiment, and you will not be told which cues to expect before each trial. You will complete 30 trials of each condition today, with a 5-minute break after 15 trials. Please complete trials as quickly and accurately as possible. When you are prompted to wait for the experimenter, I will come in to help you take off the equipment."

[Set z-file for Train "0", begin experiment, and monitor screen throughout.]

If Session 2 or 3:

"Each trial will be similar to Part II of training in that there will be 1 target 'C' and 100 distractor 'O's. Remember there is a target present in every trial. However, the order of conditions will be random throughout the experiment, and you will not be told which cues to expect before each trial. You will complete 4 blocks of 15 trials of each condition today. You may have short breaks as needed between each block. Please complete trials as quickly and accurately as possible. When you are prompted to wait for the experimenter, I will come in to help you take off the equipment."

[Confirm z-file is set for Train "0", begin experiment, and monitor screen throughout.]

Upon completion of session:

"Thank you for your participation today. Remember that you are scheduled to complete Session [2 or 3] on [insert scheduled date here]."

[Give Visa gift card(s) if appropriate.]

At the end of the data collection for the day:

[TURN OFF TACTOR CONTROL UNIT and plug it in. Switch belts for next participant if necessary.]